

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): KIMURA, et al.
Serial No.: Not yet assigned
Filed: September 17, 2003
Title: CACHE CONTROL METHOD FOR NODE APPARATUS
Group: Not yet assigned

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

September 17, 2003

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2003-075181, filed March 19, 2003.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundidge
Registration No. 29,621

CIB/alb
Attachment
(703) 312-6600

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 3月19日
Date of Application:

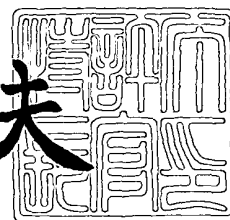
出願番号 特願2003-075181
Application Number:
[ST. 10/C]: [JP2003-075181]

出願人 株式会社日立製作所
Applicant(s):

2003年 8月26日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



出証番号 出証特2003-3069585

【書類名】 特許願

【整理番号】 PA20F787

【提出日】 平成15年 3月19日

【あて先】 特許庁長官 太田 信一郎 殿

【国際特許分類】 G06F 15/00

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

 【氏名】 木村 信二

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

 【氏名】 大島 訓

【発明者】

 【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

 【氏名】 橋本 尚

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 110000028

 【氏名又は名称】 特許業務法人 明成国際特許事務所

 【代表者】 下出 隆史

 【電話番号】 052-218-5061

【手数料の表示】

 【予納台帳番号】 133917

 【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0111082

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ノード装置におけるキャッシュ制御

【特許請求の範囲】

【請求項 1】 ストレージ装置、キャッシュ用のディスク装置を備えるノード装置、およびクライアントが接続された計算機システムにおいて、前記ディスク装置へのキャッシュを制御するキャッシュ制御方法であって、

前記ストレージ装置またはクライアントにおいて、

前記ノード装置によって中継されるデータについて、前記ディスク装置へのキャッシュの可否を表す属性情報を前記ノード装置に対して出力するステップを実行し、

前記ノード装置において、

前記属性情報に基づいて、前記中継するデータのの前記ディスク装置へのキャッシュ可否を判断するステップと、

キャッシュ否と判断されたデータについては、前記ディスク装置へのキャッシュを行わずに中継するステップとを実行するキャッシュ制御方法。

【請求項 2】 ストレージ装置、キャッシュ用のディスク装置を備えるノード装置、およびクライアントが接続された計算機システムにおいて、前記ディスク装置へのキャッシュを制御するキャッシュ制御方法であって、

前記ノード装置において、

中継すべきデータを、前記ストレージ装置またはクライアントから入力するステップと、

前記中継するデータを暗号化して前記ディスク装置に書き込むステップと、

前記ディスク装置に保持されているデータの読出コマンドに応じて、該ディスク装置から読み出したデータを復号して前記ストレージ装置またはクライアントに出力するステップとを実行するキャッシュ制御方法。

【請求項 3】 ストレージ装置、キャッシュ用のディスク装置を備えるノード装置、およびクライアントが接続された計算機システムにおいて、前記ディスク装置へのキャッシュを制御するキャッシュ制御方法であって、

前記ノード装置において、

前記ストレージ装置またはクライアント装置から受信したデータを、所定のタイミングで前記ディスク装置にキャッシュするステップを実行し、

前記ストレージ装置またはクライアントの一方において、

提供するデータについて、所定の条件に基づいて、前記ディスク装置へのキャッシュの可否を判断するステップと、

キャッシュ可と判断されたデータについては、前記ノード装置に送信するステップと、

キャッシュ否と判断されたデータについては、前記ノード装置を介さずに前記ストレージ装置またはクライアントの他方に直接データを送信するステップとを実行するキャッシュ制御方法。

【請求項4】 キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置であって、

前記中継するデータについて、前記ディスク装置へのキャッシュの可否を表す属性情報を入力する属性情報入力部と、

該属性情報に基づいて、前記中継するデータの前記ディスク装置へのキャッシュ可否を判断する判断部と、

キャッシュ否と判断されたデータについては、前記ディスク装置へのキャッシュを行わずに中継するキャッシュ制御部とを備えるノード装置。

【請求項5】 請求項4記載のノード装置であって、

前記ディスク装置の他に、キャッシュ用の揮発性メモリを有し、

前記キャッシュ制御部は、

前記キャッシュ可否に関わらず前記データを前記揮発性メモリにキャッシュするメモリ制御部と、

前記揮発性メモリから前記ディスク装置への書き込みを行うための所定の条件が成立した場合に、前記キャッシュ否と判断されたデータを除き、該揮発性メモリに保持されたデータを前記ディスク装置に移転する移転制御部とを備えるノード装置。

【請求項6】 キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置であって、

前記中継するデータを暗号化して前記ディスク装置に書き込む暗号化部と、
前記ディスク装置に保持されているデータの読出コマンドに応じて、該ディスク装置から読み出したデータを復号して前記ストレージ装置またはクライアントに出力する復号部とを備えるノード装置。

【請求項 7】 請求項 6 記載のノード装置であって、
前記中継するデータについて、暗号化の要否を表す属性情報を入力する属性情報入力部と、

該属性情報に基づいて、前記中継するデータの暗号化の要否を判断する判断部と、

暗号化要と判断されたデータについては、前記暗号化部および復号部を用いて前記ディスク装置へのキャッシュを行うキャッシュ制御部とを備えるノード装置。

【請求項 8】 請求項 6 記載のノード装置であって、
前記ストレージ装置から前記暗号化および復号に使用する鍵データを受信し、揮発性メモリに管理する鍵データ管理部を備えるノード装置。

【請求項 9】 請求項 4～8 いずれか記載のノード装置であって、
前記属性情報は、前記データの中継時に、該データと関連付けて入力されるノード装置。

【請求項 10】 請求項 4～8 いずれか記載のノード装置であって、
前記属性情報入力部は、前記ストレージ装置におけるデータの格納ブロックと前記属性情報との関係を予め取得し、管理するノード装置。

【請求項 11】 請求項 4～8 いずれか記載のノード装置であって、
前記属性情報は、前記クライアントにおける異常発生を通知する情報であるノード装置。

【請求項 12】 ディスク装置を有するノード装置を介して、クライアントに、データを提供するストレージ装置であって、
前記データについて、前記ディスク装置へのキャッシュの可否、または前記ディスク装置への書き込み時の暗号化の要否を表す属性情報を管理する属性情報管理部と、

前記ノード装置に、該属性情報を通知する属性情報通知部とを備えるストレージ装置。

【請求項 13】 請求項 12 記載のストレージ装置であって、
前記属性情報通知部は、前記データの提供時に、該データと関連づけて前記属性情報を通知するストレージ装置。

【請求項 14】 請求項 12 記載のストレージ装置であって、
前記属性情報通知部は、前記データの格納ブロックと前記属性情報との関係をデータの提供に先立って前記ノード装置に通知するストレージ装置。

【請求項 15】 請求項 12 記載のストレージ装置であって、
前記ディスクへの書き込み時にデータを暗号化可能な複数のノード装置と接続されており、
いずれかの前記ノード装置に対して前記暗号化に用いる鍵データの出力を行うべき所定の条件が満たされたか否かを判断する条件判断部と、
該ノード装置に対して前記鍵データを送信する鍵データ送信部とを備えるストレージ装置。

【請求項 16】 ディスク装置を有するノード装置を介してストレージ装置にデータを書き込み可能なクライアントであって、
前記ディスク装置へのキャッシュの禁止、または前記ディスク装置への書き込み時の暗号化を前記ノード装置に要求するための所定の条件が、満たされたか否かを判断する条件判断部と、

該条件が満たされた時に、前記ノード装置に対して、前記要求を表す属性情報を出力する禁止通知部と、

該属性情報の出力と併せて、または該属性情報の出力後に、前記データをノード装置に送信するデータ送信部とを備えるクライアント。

【請求項 17】 請求項 16 記載のクライアントであって、
前記所定の条件は、該クライアントにおけるソフトウェアの動作異常時であるクライアント。

【請求項 18】 キャッシュ用のディスク装置を備えるノード装置を介して、他の計算機にデータを提供するデータ提供装置であって、

前記提供するデータについて、所定の条件に基づいて、前記ディスク装置へのキャッシュの可否を判断する判断部と、

キャッシュ可と判断されたデータについては前記ノード装置に送信し、キャッシュ否と判断されたデータについては、前記ノード装置を介さずに前記他の計算機に送信する送信制御部とを備えるデータ提供装置。

【請求項 19】 請求項 18 記載のデータ提供装置であって、

前記提供するデータについて、キャッシュの可否を表す属性情報を予め管理する属性情報管理部を備え、

前記判断部は、該属性情報に基づいてキャッシュの可否を判断するデータ提供装置。

【請求項 20】 請求項 18 記載のデータ提供装置であって、

前記判断部は、該データ提供装置におけるソフトウェアの動作異常時に、キャッシュ否と判断するデータ提供装置。

【請求項 21】 請求項 18 記載のデータ提供装置であって、

前記他の計算機にデータを送信するためのアドレス情報を管理するアドレス管理部を備え、

前記送信制御部は、前記キャッシュ否と判断された場合には、データの送信先を前記ノード装置のアドレスから前記アドレス管理部で管理されているアドレス情報に切り換えるデータ提供装置。

【請求項 22】 ストレージ装置と複数のクライアントとの間で、データの授受を中継するノード装置であって、

いずれかの前記クライアントから前記ストレージ装置内のデータの読出コマンドを受け付ける要求受付部と、

該読出コマンドを発信した前記クライアントを特定可能な情報を、該読出コマンドとともに、前記ストレージ装置に送信する要求送信部とを備えるノード装置

。

【請求項 23】 キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置の動作を制御するコンピュータプログラムであって、

前記中継するデータについて、キャッシュの可否を表す属性情報を入力する機能と、

該属性情報に基づいて、前記中継するデータのキャッシュ可否を判断する機能と、

キャッシュ否と判断されたデータについては、前記ディスク装置へのキャッシュを行わずに中継する機能とを該ノード装置において実現するためのコンピュータプログラム。

【請求項 24】 キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置の動作を制御するためのコンピュータプログラムであって、

前記中継するデータを暗号化して前記ディスク装置に書き込む機能と、

該ディスク装置から読み出したデータを復号して前記ストレージ装置またはクライアントに出力する機能とを該ノード装置において実現するためのコンピュータプログラム。

【請求項 25】 ディスク装置を有するノード装置を介して、クライアントに、データを提供するストレージ装置の動作を制御するためのコンピュータプログラムであって、

前記データについて、前記ディスク装置へのキャッシュの可否、または前記ディスク装置への書き込み時の暗号化の可否を表す属性情報を管理する機能と、

前記ノード装置に、該属性情報を通知する機能とを該ストレージ装置において実現するためのコンピュータプログラム。

【請求項 26】 ディスク装置を有するノード装置を介してストレージ装置にデータを書き込み可能なクライアントの動作を制御するためのコンピュータプログラムであって、

前記ディスク装置へのキャッシュの禁止、または前記ディスク装置への書き込み時の暗号化を、前記ノード装置に要求するための所定の条件が満たされたか否かを判断する機能と、

該条件が満たされた時に、前記ノード装置に対して、前記要求を表す属性情報を出力する機能と、

該属性情報の出力と併せて、または該属性情報の出力後に、前記データをノード装置に送信する機能とを該クライアントにおいて実現するためのコンピュータプログラム。

【請求項 27】 キャッシュ用のディスク装置を備えるノード装置を介して、他の計算機にデータを提供するデータ提供装置の動作を制御するコンピュータプログラムであって、

前記提供するデータについて、所定の条件に基づいて、前記ディスク装置へのキャッシュの可否を判断する機能と、

キャッシュ否と判断されたデータについては、前記ノード装置を介さずに前記他の計算機に直接データを送信する機能とを該データ提供装置において実現するためのコンピュータプログラム。

【請求項 28】 ストレージ装置と複数のクライアントとの間で、データの授受を中継するノード装置の動作を制御するコンピュータプログラムであって、

いずれかの前記クライアントから前記ストレージ装置内のデータの読出コマンドを受け付ける機能と、

該読出コマンドを発信した前記クライアントを特定可能な情報を、該読出コマンドとともに、前記ストレージ装置に送信する機能とを該ノード装置において実現するためのコンピュータプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ストレージ装置、クライアント間でデータを中継するノード装置におけるキャッシュの制御に関する。

【0002】

【従来の技術】

企業などでは、複数のクライアントコンピュータ（以下、単にクライアントと称する）をネットワークで接続した大規模な計算機システムが構築されている。このような計算機システムでは、各クライアントで業務を遂行する際に参照する種々のデータやプログラムを供給するため、ストレージ装置と呼ばれる大容量の

記憶装置が設けられている。ストレージ装置は、大容量のハードディスクを備えており、各クライアントのユーザは、自己に割り当てられた領域、または共有の領域に対して、データやプログラムなどの格納・読み出しを行う。通常、ストレージ装置へのアクセスには、SCSI (Small Computer System Interface) と呼ばれるプロトコルが使用される。近年では、SCSI用のコマンドを、いわゆるIPパケットに分割してネットワーク上を転送するためのiSCSIと呼ばれるプロトコルが提案されている。iSCSIを用いることにより、従来よりも更に多数のクライアント、および遠隔地のクライアントがストレージ装置にアクセス可能な環境が提供されつつある。

【0003】

ストレージ装置の活用方法の一つとして、リモートブートと呼ばれるクライアントの起動方法が提案されている。リモートブートでは、起動に必要なオペレーティングシステム、環境設定ファイルなどの各種ファイルが、予めストレージ装置の各ユーザ用の領域に格納されている。各クライアントは、ネットワークを介してこれらのファイルをストレージ装置から読み込むことにより、起動することができる。

【0004】

リモートブートする場合、クライアントは、ストレージ装置へのアクセスを実現するための比較的小さなブートプログラムをROMに記憶していれば十分であり、ハードディスクを備えている必要がない。ハードディスクに格納されるべきデータはストレージ装置で一元管理されるから、仮にクライアントコンピュータが破損、盗難などに遭っても、重要なデータの遺失、漏洩を逃れることができるというセキュリティ上の利点がある。また、各ユーザは、ネットワーク上に接続されたいずれのクライアントも、一定の環境設定で起動させることができるという実用上の利点もある。リモートブートについては、例えば、特許文献1～3に開示されている。

【0005】

ストレージ装置を用いる場合、多数のクライアントによるストレージ装置へのアクセスの集中を緩和するため、クライアントとストレージ装置との間に、キャ

ッシュ用のメモリを備えたノード装置を設けることがある。かかるシステムでは、ノード装置にキャッシュ済みのデータは、ストレージ装置にアクセスするまでなくクライアントに供給可能であるため、ストレージ装置へのアクセスが緩和される。

【0006】

【特許文献1】

特開平6-332716号公報

【特許文献2】

特開2001-75853号公報

【特許文献3】

特開2000-259583号公報

【0007】

【発明が解決しようとする課題】

しかし、ノード装置を用いたシステムでは、このノード装置がセキュリティ上、脆弱になり易いという課題があった。一般に、ストレージ装置は、企業の機密データを一元的に保持しているため、非常に厳しくセキュリティ管理されている。例えば、ストレージ装置が設置された部屋への立ち入りは厳しく制限されており、ソフトウェア的にもファイアウォールなどでストレージ装置へのアクセスは厳しく制限されている。一方、クライアントは、ハードディスクを備えないディスクレスコンピュータとすることにより、データの漏洩を防止することができる。仮に、クライアントがハードディスクを備えている場合でも、各ユーザは、自己が使用する装置については盗難や不正使用に遭わないよう気を配るため、一定レベルでセキュリティは確保される。

【0008】

これに対し、ノード装置に対しては、ストレージ装置ほどセキュリティが確保された環境が確保されておらず、各ユーザもセキュリティに気を配らないことが多い。かかる環境下で、ノード装置には、ストレージ装置とクライアントの間で授受される機密データがキャッシュされている。従って、ノード装置の盗難や不正アクセスによって、機密データが漏洩する可能性があった。かかる機密データ

の一つとして、例えば、動作に支障が生じた時にクライアントが出力するコアダンプが挙げられる。コアダンプは、通常のコンピュータでは、内蔵のハードディスクに書き出されるが、ディスクレスコンピュータではノード装置にキャッシュされた後、ストレージ装置に書き出される。コアダンプには、ストレージ装置にアクセスするためのパスワードなど、コンピュータに保存されていた種々の機密情報が含まれており、これらの情報の漏洩は、ストレージ装置への不正アクセスを許容する原因となり得る。本発明は、これらの課題に鑑み、ノード装置におけるセキュリティの向上を図ることを目的とする。

【0009】

【課題を解決するための手段】

本発明では、ストレージ装置、ノード装置、およびクライアントが接続された計算機システムにおいて、ノード装置に備えられたキャッシュ用のディスク装置へのキャッシュを次の方法で制御する。この制御は、ノード装置の揮発性メモリへのキャッシュにも同様に適用してもよいし、ディスク装置へのキャッシュ時のみ適用してもよい。

【0010】

＜第1の制御方法＞ 第1の制御方法では、ディスク装置へのキャッシュの可否を表す属性情報を用いる。ストレージ装置またはクライアントは、ノード装置によって中継されるデータについて、上述の属性情報をノード装置に対して出力する。ノード装置は、この属性情報に基づいて、キャッシュ否と判断されたデータについては、ディスク装置へのキャッシュを行わずに中継する。こうすることにより、機密データについてはノード装置でのキャッシュを禁止することができ、セキュリティを確保することができる。

【0011】

ノード装置がディスク装置の他に、キャッシュ用の揮発性メモリを有している場合、キャッシュ否のデータについては揮発性メモリへのキャッシュを禁止してもよい。また、揮発性メモリに対しては、キャッシュ可否に関わらずキャッシュを行い、揮発性メモリからディスク装置への書き込みは、キャッシュ否と判断されたデータを除いて行うようにしてもよい。

【0012】

キャッシュ可否を制御するための属性情報は、種々のタイミングで取得することができる。例えば、データの中継時に、このデータと関連付けて属性情報を入力するようにしてもよい。ヘッダ等に属性情報を含めることでデータと属性情報を一体として入力してもよいし、中継されるデータと分けて属性情報のみを入力してもよい。かかる態様は、例えば、ストレージ装置において、データごとに属性情報を予め管理しておき、ノード装置に、この属性情報を通知する機能をもたせることにより、実現することができる。

【0013】

ストレージ装置において、データの格納ブロックと対応づけて属性情報が管理されている場合、ノード装置は、この対応を予め取得し、管理しておいてもよい。こうすれば、データの中継時にストレージ装置から属性情報を受け取るまでなく、キャッシュの可否を判断することができる。

【0014】

属性情報は、必ずしもデータ単位でキャッシュの可否を通知する情報である必要はない。例えば、クライアントにおけるソフトウェアの動作異常発生を通知する情報を属性情報として用いても良い。ノード装置は、この通知を受信した後は、そのクライアントからのデータのキャッシュを行わないというように、キャッシュ可否の動作を切り替えることができる。同様に、クライアントの動作が正常に復帰したことを通知する情報を属性情報として用いても良い。このように、属性情報は、ノード装置におけるキャッシュ可否を切り替えるタイミングを通知する情報とすることもできる。

【0015】

かかる通知は、キャッシュの禁止をノード装置に要求するための所定の条件が満たされた時に、ノード装置に対して属性情報を出力する機能をクライアントに持たせることにより実現できる。クライアントは、この属性情報の出力と併せて、または該属性情報の出力後に、データをノード装置に送信するようにすればよい。キャッシュの禁止を要求するための条件は、例えば、上述の動作異常の他、予め設定された特定のアプリケーションの起動時など種々の設定が可能である。

【0016】

＜第2の制御方法＞ 本発明の第2の制御方法では、ノード装置において、中継すべきデータを暗号化してディスク装置に書き込むことでキャッシュする。キャッシュされたデータの読出コマンドを受けた時には、ディスク装置から読み出したデータを復号してストレージ装置またはクライアントに出力する。暗号化は、ディスク装置への書き出し時に完了していればよく、例えば、書き出し直前に行ってもよいし、データの受信時に行ってもよい。こうすることで、ディスク装置が盗難に遭っても、データのセキュリティを確保することができる。

【0017】

第2の制御方法においては、必ずしも全てのデータを暗号化する必要はない。例えば、中継するデータについて、暗号化の要否を表す属性情報を入力し、この属性情報に基づいて、暗号化の要否を切り替えても良い。属性情報は、第1の制御方法でキャッシュの可否を制御するために例示した種々の情報を適用することができる。こうすれば、不要な暗号化を抑制することができ、ノード装置での処理負担を軽減することができる。第1の制御方法と同様、属性情報は、暗号化の要否を切り替えるタイミングを通知する情報としてもよい。

【0018】

第2の制御方法においては、暗号化および復号に用いる鍵データを揮発性メモリに管理することが好ましい。こうすれば、ノード装置が盗難等にあった場合には、鍵データが消失するため、ディスク装置に記憶されたデータの復号が不能となる。この鍵データは、例えば、ストレージ装置がノード装置に与えるようにしてもよいし、ノード装置自身が生成するようにしてもよい。ストレージ装置から鍵データを供給する場合、ストレージ装置は、各ノード装置に与えられる鍵データに重複が生じないように管理することが好ましい。こうすることにより、与えられた鍵データは他のノード装置には適用できなくなるため、ノード装置へのなりすましによって鍵データが取得された場合でも、セキュリティを確保することが可能となる。ストレージ装置からノード装置への鍵データの供給は、例えば、ノード装置が新たに稼働し始めた場合や、ノード装置の管理者がストレージ装置に対して鍵データの発行を要求した場合など、種々の条件下で行うことができる。

【0019】

＜第3の制御方法＞ 第3の制御方法は、ストレージ装置およびクライアントの一方から他方へのデータ送信時において、データの発信側となるデータ提供装置が、データの送信先を切り換えることにより実現される。データ提供装置は、提供するデータについて、ディスク装置へのキャッシュの可否を判断する。キャッシュ可否の判断条件は、例えば、予めデータごとに設定しておいてもよいし、「ソフトウェアの動作異常時」など特定の事象やタイミングとの関係で設定してもよい。データ提供装置はキャッシュ可と判断されたデータについては、ノード装置に送信し、ノード装置は、このデータをディスク装置にキャッシュする。キャッシュ否と判断されたデータについては、ノード装置を介さずに、受信側の装置に、直接データを送信する。こうすることにより、ノード装置では特別な制御を要することなく、キャッシュを回避することができる。

【 0 0 2 0 】

第3の制御方法を実現するため、データ提供装置は、ノード装置とは別に、受信側となる計算機のアドレス情報を予め管理しておいてもよい。データ提供装置は、キャッシュ否と判断された場合には、データの送信先をノード装置のアドレスから、予め管理されているアドレス情報に切り換えることにより、ノード装置を迂回した通信を実現することができる。この態様は、受信側となる計算機のアドレス情報が予め固定されている場合、例えば、クライアントからストレージ装置へのデータ送信時に有用である。

【 0 0 2 1 】

別の態様として、ノード装置は、ストレージ装置内のデータの読出コマンドを受け付けた時、この読出コマンドを発信したクライアントを特定可能な情報、例えば、アドレス情報、クライアント名、クライアントの識別情報などを、ストレージ装置に送信するようにしてもよい。こうすることでストレージ装置は、ノード装置を迂回する場合のデータの送信先を特定することができる。ストレージ装置が、ノード装置を迂回してクライアントにデータを送信する場合、ノード装置に対して、データの送信が完了した通知を行うようにしてもよい。この通知がノード装置からストレージ装置に出力した読出コマンドに対する応答となり、ノード

ド装置を応答待ちの状態から解放することができる。

【0022】

本発明は、上述したノード装置、ストレージ装置、クライアントに限らず種々の態様で構成することが可能である。例えば、これらを接続した計算機システムとして構成してもよい。また、ノード装置におけるキャッシュを制御する制御方法として構成してもよい。コンピュータによりこの制御を実現するためのコンピュータプログラムおよびこれを記録した記憶媒体として構成してもよい。ここで、記憶媒体としては、フレキシブルディスクやCD-ROM、DVD、光磁気ディスク、ICカード、ROMカートリッジ、パンチカード、バーコードなどの符号が印刷された印刷物、コンピュータの内部記憶装置（RAMやROMなどのメモリ）および外部記憶装置などコンピュータが読取り可能な種々の媒体を利用できる。

【0023】

【発明の実施の形態】

本発明の実施の形態について以下の順序で説明する。

A. 第1実施例；

- A1. システム構成；
- A2. 機能ブロック；
- A3. テーブルのデータ構造；
- A4. データ読出処理；
- A5. データ書込処理；
- A6. 変形例；

B. 第2実施例；

- B1. 機能ブロック；
- B2. データ読出処理；
- B3. データ書込処理；

C. 第3実施例；

- C1. 機能ブロック；
- C2. データ読出処理；

C3. データ書込処理;

【0024】

A. 第1実施例;

A1. システム構成;

図1は第1実施例としての計算機システムの構成を示す説明図である。この計算機システムでは、ストレージ装置300に、複数のノード装置200および計算機100がIPパケットを授受可能なネットワークで接続されている。ネットワークとしては、LAN (Local Area Network)、イントラネット、インターネットなどを適用することができる。

【0025】

ストレージ装置300は、図示するハードウェア構成を有するコンピュータとして構成されており、大容量のディスク装置310に蓄積されたデータをノード装置200経由で各計算機100に提供することができる。また、各計算機100はハードディスク310にデータを格納することもできる。NIC (Network Interface Card) 304は、ネットワークを介してデータを授受するためのインタフェースである。I/O 305は、入出力装置用のインタフェースである。

【0026】

ストレージ装置300では、CPU 301が、ディスク装置310およびROM 303に記憶された制御用のプログラムに従って、データの授受を制御する。メモリ302は、このプログラムの動作時に使用される主記憶である。

【0027】

ノード装置200は、図示するハードウェア構成を有するコンピュータとして構成されており、ストレージ装置300と計算機100の間で授受されるデータを中継する機能を奏する。ディスク装置210およびメモリ202は、中継されるデータを一時的に蓄積するキャッシュのために利用される。NIC (Network Interface Card) 204は、ネットワークを介してデータを授受するためのインタフェースである。I/O 205は、入出力装置用のインタフェースである。

【0028】

ノード装置200では、CPU 201が、ディスク装置210およびROM 2

03に記憶された制御用のプログラムに従って、データの中継およびキャッシュを制御する。メモリ202は、キャッシュの他、このプログラムの動作時にも使用される。本実施例では、中継するデータによってキャッシュの可否が切り替えられる。キャッシュが禁止されているデータについては、ノード装置200は、キャッシュを行わずに中継する。キャッシュの禁止については、二通りの制御が適用可能である。第1の態様は、キャッシュが禁止されているデータに関しては、ノード装置200は一切、キャッシュを行わないという制御方法である。第2の態様は、ノード装置200は揮発性のメモリ202には、キャッシュ可否の属性に関わらずキャッシュを行い、メモリ202からディスク装置210にデータを移転する際に、キャッシュの可否を判断する制御方法である。本実施例では、第1の態様でキャッシュの可否を制御するものとして説明し、後で変形例として第2の態様について説明する。

【0029】

計算機100は、ハードディスクを備えないディスクレスコンピュータである。内部には、CPU101、RAM102、ROM103、NIC104およびI/O105が備えられている。計算機100は、起動時には、必要なオペレーティングシステムなどのファイルをストレージ装置300から読み込むことで、リモートブートする。ROM103には、リモートブート用のプログラムが予め記憶されている。

【0030】

ストレージ装置300、ノード装置200、計算機100の動作を制御するためのプログラムは、CD-ROMなどの記録媒体によってストレージ装置に供給してもよい。

【0031】

A2. 機能ブロック;

図2は計算機システムを構成する各装置内の機能ブロックを示す説明図である。本実施例では、図示する各機能ブロックは、それぞれコンピュータプログラムをインストールすることによりソフトウェア的に構成される。各機能ブロックは、ハードウェア的に構成することも可能である。

【0032】

計算機100の機能ブロックは、次の通りである。TCP/IPモジュール120は、ネットワークを介してIPパケットを通信する。SCSIモジュール122は、ストレージ装置300に対してデータの読み出し、書き込み用のコマンド送受信などを行う。iSCSIモジュール121は、SCSIコマンドとIPパケットの変換を行うことによって、SCSIコマンドをTCP/IPで送受信可能とする。ブートプログラム130は、計算機100でリモートブートを実現する。ブートプログラム130は、計算機100の初期化を行う機能、起動に必要なオペレーティングシステムなどのファイルをストレージ装置から読み込む機能、これらのファイルを起動してブートする機能などを実現することができる。ID入力モジュール131は、リモートブートを行う際に、ストレージ装置300に送信すべきユーザID、パスワードを入力する。本実施例では、これらのモジュールは、ROM103に記憶されているプログラムによって実現される。

【0033】

上述の各機能ブロックによってリモートブートが完了すると、計算機100には、ストレージ装置300から供給されたファイルによってオペレーティングシステム140が起動する。オペレーティングシステム140には、計算機100の動作に異常が生じた時に、種々の情報をコアダンプとしてストレージ装置300に書き出す機能を奏する異常制御モジュール141が備えられている。計算機100上の各処理に利用されるアプリケーション150は、ストレージ装置300から供給され、オペレーティングシステム140の上で稼働する。

【0034】

本実施例では、クライアントごとに仮想ボリュームが与えられる。仮想ボリュームとは、ストレージ装置300のディスク装置310内に設けられた物理的な領域ではなく、各クライアントのデータを管理するための論理的なディスクを意味する。仮想ボリュームは、ユーザIDに対応づけて設定されており、ユーザが、計算機100でユーザIDを入力することによって、計算機100と対応づけられる。以下、本実施例では、「クライアント」という用語は、単にハードウェアとしての計算機ではなく、ユーザがログインし、仮想ボリュームと対応づけら

れた状態にある計算機を意味するものとする。

【0035】

ノード装置 200 には、ネットワークを介して SCS I コマンドを授受するため、TCP/IP モジュール 220、iSCSI モジュール 221、SCSI モジュール 222 が備えられている。ノード装置 200 には、所定のオペレーティングシステムがインストールされており、このオペレーティングシステム上で稼働するアプリケーションによってキャッシュ制御モジュール 230 が実現されている。

【0036】

キャッシュ制御モジュール 230 は、キャッシュ属性管理部 234 およびキャッシュ管理テーブル 232 を参照して、ノード装置 200 におけるキャッシュを制御する。キャッシュ管理テーブル 232 とは、クライアントから指定された仮想ブロックと、ノード装置 200 におけるデータの格納場所、即ちキャッシュブロックとを対応づけるテーブルである。キャッシュ属性管理部 234 とは、仮想ブロックの各データについて、キャッシュの可否を示す属性情報を管理するテーブルである。これらのテーブルのデータ構造については後述する。

【0037】

ストレージ装置 300 には、ネットワークを介して SCS I コマンドを授受するため、TCP/IP モジュール 320、iSCSI モジュール 321、SCSI モジュール 322 が備えられている。ストレージ装置 300 には、所定のオペレーティングシステムがインストールされており、このオペレーティングシステム上で稼働するアプリケーションによってアクセス制御モジュール 330 が実現されている。

【0038】

アクセス制御モジュール 330 は、ユーザ管理テーブル 334、物理ブロック管理テーブル 332 を参照して、ディスク装置 310 のデータの読み書きを制御する。ユーザ管理テーブル 334 は、ユーザ ID と仮想ボリュームとを対応づけるテーブルである。物理ブロック管理テーブル 332 とは、仮想ブロックと物理ブロックとを対応づけるテーブルである。アクセス制御モジュール 330 は、計

算機 100 の起動時に、計算機 100 からユーザ ID を取得する。このユーザ ID に基づいてユーザ管理テーブル 334 を参照することで、ユーザに対応する仮想ボリュームを計算機 100 に割り当てる。アクセス制御モジュール 330 は、以後、クライアントから指定された仮想ブロックに対して、物理ブロック管理テーブル 332 を参照して、ディスク装置 310 のデータの読み書きを制御することができる。

【0039】

キャッシュ管理テーブル 333 は、仮想ブロックごとにキャッシュの可否を記録したテーブルである。アクセス制御モジュール 330 は、キャッシュ属性管理テーブル 333 の内容をノード装置 200 に通知することで、ノード装置 200 におけるキャッシュの可否を制御する。

【0040】

A3. テーブルのデータ構造；

図 3 はユーザ管理テーブル 334 の構造を示す説明図である。ユーザ管理テーブル 334 は、ユーザ ID、パスワード、仮想ボリューム、属性、IP アドレスを対応づけて管理する。属性とは、仮想ボリュームに対する他のユーザのアクセスの可否を表している。「専有」は他のユーザのアクセスが禁止されていることを意味し、「共有」は許可されていることを意味する。IP アドレスは、ユーザがログインした計算機によって変動する。属性、IP アドレスは、ユーザ管理テーブル 334 から省略してもよい。

【0041】

図 4 は物理ブロック管理テーブル 332 の構造を示す説明図である。図の上方にデータ構造を示し、下方に仮想ブロックと物理ブロックとの対応関係を模式的に示した。物理ブロック管理テーブル 332 は、仮想ブロックと物理ブロックの対応関係を示すテーブルであり、仮想ボリューム単位で設けられている。図中には、仮想ボリューム VDa に対応するテーブルを例示した。このテーブルでは、仮想ボリューム VDa の仮想ブロック BL0 ～ BLn に対して、実際にデータが格納されている物理ブロックが記憶されている。物理ブロックは、物理ボリュームとブロック番号との組み合わせで定義される。

【0042】

図に例示したテーブルは、下方に模式的に示した対応関係を表している。つまり、仮想ブロックBL0のデータは、物理ボリュームPD11のブロックBL0に格納されている。同様にして、仮想ブロックBL1～BLnの各データは、それぞれ物理ボリュームPD11、PD12のいずれかのブロックに格納されている。図では、ディスク装置310が、パーティションによって複数の物理ボリュームに分割されている場合を例示したが、単一の物理ボリュームとして構成されていてもよい。この場合、物理ブロック管理テーブル332からは、物理ボリュームに関するデータを省略しても構わない。

【0043】

図5はキャッシュ管理テーブル232の構造を示す説明図である。このテーブルは、仮想ブロックと、ノード装置200におけるキャッシュブロックとを対応づけている。ディスク書込の欄については、後述する。ノード装置200におけるキャッシュの格納場所には、揮発性のメモリ202と、ディスク装置210の2種類が存在する。従って、本実施例では、格納場所とブロック番号で、キャッシュブロックを特定するものとした。

【0044】

図6はキャッシュ属性管理テーブル333の構造を示す説明図である。キャッシュ属性管理テーブル333は、データの読み出しと書き込みに分けて用意されており、仮想ボリュームごとにキャッシュの可否を表す情報を格納している。本実施例では、ディスク装置が複数の物理ボリュームに分割されているため、物理ボリューム単位でキャッシュの可否を設定可能とした。例えば、図の例では、仮想ボリュームVDaの読み出しについて、物理ボリュームPD11に格納されたデータについてはキャッシュ可、物理ボリュームPD12に格納されたデータについてはキャッシュ不可と設定されている。

【0045】

図の例において、物理ボリュームPD12は、仮想ボリュームVDa、VDbに共有されている。本実施例では、このように複数の仮想ボリュームによる物理ボリュームの共有を許容してもよい。この場合、共有される物理ボリュームに格

納されたデータは、各仮想ボリュームに対応した複数のクライアントで共有されることになる。図中の例では、共有されている物理ボリュームPD12について、キャッシュ属性は不可と設定されている場合を例示した。これに対し、仮想ボリュームVDa、VDbで、共有されている物理ボリュームPD12に対するキャッシュ属性を異ならせても良い。こうすることにより、例えば、仮想ボリュームVDaからの読み出し時にはキャッシュを許可し、仮想ボリュームVDbからの読み出し時にはキャッシュを禁止するというように、読み出すクライアントに応じてキャッシュの可否を切り換えることが可能となる。

【0046】

書き込みについても、同様の形式でキャッシュの可否が管理されている。キャッシュ可否については、計算機システムの動作時に動的に設定を変更してもよい。例えば、仮想ボリュームVDaを使用するクライアントの動作異常を検知した時、ノード装置200は仮想ボリュームVDaに対応する全物理ボリュームについて、一時的に「キャッシュ不可」の属性を設定するようにしてもよい（図中のデータ領域Da参照）。

【0047】

A4. データ読出処理；

図7はデータ読み出し処理のフローチャートである。リモートブート時を例にとり、左側にクライアント100の処理、中央にノード装置200の処理、右側にストレージ装置300の処理を示した。

【0048】

ユーザが計算機の電源をオンにすると、ID入力モジュール131によって、ユーザID、パスワード等の入力画面が提示される。ユーザが、ユーザID、パスワードを入力し、ログインすると、クライアント100は、IPアドレスとともにこれらの情報を、ノード装置200およびストレージ装置300に送信する（ステップS10）。ストレージ装置300は、この情報に基づいてユーザ管理テーブル334を参照し、ユーザの認証を行う（ステップS30）。正規なユーザであることが確認されると、ストレージ装置300は、ユーザ管理テーブル334に、クライアント100のIPアドレスを記録するとともに、ユーザに対応

した仮想ボリューム名を通知する。この通知は、ノード装置 200 を介してクライアント 100 に送信される。以後、クライアント 100 からのデータの読み出し要求等は、この仮想ボリューム名に基づいて行われる。

【0049】

クライアント 100 は、ROM に記録されているブートプログラム 130 を起動する（ステップ S11）。ブートプログラム 130 は、ブートに必要な初期化処理を行った上で、オペレーティングシステムのデータ読み出し要求をノード装置に対して送信する（ステップ S12）。

【0050】

ノード装置 200 は、読み出し要求を受信し、キャッシュ管理テーブル 232 を参照して、要求されたデータが、既にキャッシュ済みであるか否かを確認する（ステップ S20）。キャッシュ済みの場合は（ステップ S20）、キャッシュブロックから該当するデータを読み出し（ステップ S26）、クライアント 100 に返信する（ステップ S25）。クライアント 100 は、このデータを受信して起動処理を継続する（ステップ S13）。

【0051】

要求されたデータがキャッシュされていない（以下、「未キャッシュ」と呼ぶ）場合（ステップ S20）、ノード装置 200 はストレージ装置 300 にデータの読み出しを要求する（ステップ S21）。ストレージ装置 300 は、この要求に応じて、物理ブロック管理テーブル 332 を参照して、物理ボリュームからデータを読み出し、ノード装置 200 に返信する（ステップ S32）。これと併せて、ストレージ装置 300 は、キャッシュ属性管理テーブル 333 を参照し、要求されたデータのキャッシュ可否の属性情報をノード装置 200 に通知する（ステップ S32）。この属性情報は、要求されたデータのヘッダ等に埋め込んだ形で通知してもよいし、要求されたデータと分けて通知してもよい。

【0052】

ノード装置 200 は、ストレージ装置 300 からデータおよび属性情報を受信すると、属性情報に基づいてキャッシュの可否を判断する（ステップ S22）。キャッシュ可と判断される場合には、受信したデータをキャッシュし（ステップ

S 2 3)、その格納先に応じてキャッシュ管理テーブル 2 3 2 を更新して (ステップ S 2 4)、データをクライアントに返信する (ステップ S 2 5)。キャッシュ不可と判断される場合には、上述のキャッシュ処理をスキップし、データをクライアントに返信する (ステップ S 2 5)。クライアントはこうしてノード装置 2 0 0 から送信されたデータを受信する (ステップ S 1 3)。

【0053】

以上の処理を繰り返し行うことにより、クライアント 1 0 0 は、ブートに必要なファイルを取得し、リモートブートを完了することができる。図 7 ではリモートブート時の読み出し処理を示したが、リモートブートと無関係に通常データの読み出し時にも、データ読み出し要求 (ステップ S 1 2) からデータ受信 (ステップ S 1 3) の処理が同様に行われる。

【0054】

上述の実施例では、データの要求時にノード装置 2 0 0 からストレージ装置 3 0 0 にキャッシュの可否を問い合わせる場合を例示した。これに対し、ノード装置 2 0 0 は、ストレージ装置 3 0 0 から、予めキャッシュ属性管理テーブル 3 3 3 の内容をまとめて取得し、キャッシュ属性管理部 2 3 4 で管理するようにしてもよい。こうすることにより、データ要求のたびに、属性情報の問い合わせを行う必要性をなくすことができる。

【0055】

A 5. データ書込処理;

図 8 はデータ書込処理のフローチャートである。左側にクライアント 1 0 0 の処理、中央にノード装置 2 0 0 の処理、右側にストレージ装置 3 0 0 の処理を示した。クライアント 1 0 0 は既に起動し、ノード装置 2 0 0、ストレージ装置 3 0 0 とデータ授受が可能な状態で、クライアント 1 0 0 の動作にソフトウェア上の異常が生じた場合のコアダンプの書き込みを例示した。

【0056】

クライアント 1 0 0 の異常制御モジュール 1 4 1 は、ソフトウェア上の動作異常を検出すると (ステップ S 4 0)、異常の発生を IP アドレスとともにノード装置 2 0 0 に通知する (ステップ S 4 0)。ノード装置 2 0 0 は、この情報に基

づいてキャッシュ属性管理部 234 における属性情報を更新する（ステップ S 50）。先に図 6 で示したように、このクライアントからの書き込みについては、全てキャッシュ不可という属性に変更するのである。

【0057】

クライアント 100 の異常制御モジュール 141 は、仮想ブロックを指定して、コアダンプを書き込むための要求をノード装置に送信する（ステップ S 42）。ノード装置 200 は、書き込み要求を受信し、キャッシュ属性管理部 234 を参照して、キャッシュの可否を判断する（ステップ S 52）。クライアント 100 の動作に異常が生じた場合には、ノード装置 200 は、ステップ S 50 で更新された属性情報に基づき、キャッシュ不可と判断することになる。一方、クライアント 100 に異常が生じる前、異常から回復した後、および別のクライアントからの書き込み要求時などには、指定された仮想ブロックの属性情報に基づき、ノード装置 200 はキャッシュ可と判断する場合もある。

【0058】

キャッシュ可と判断した場合（ステップ S 52）、ノード装置は、クライアントから受信したデータをキャッシュし（ステップ S 53）、その格納先に応じてキャッシュ管理テーブル 232 を更新するとともに（ステップ S 54）、データをストレージ装置 300 に送信する（ステップ S 55）。キャッシュ不可と判断した場合（ステップ S 52）、ノード装置 200 は、クライアントから受信したデータをキャッシュすることなく、ストレージ装置 300 に送信する（ステップ S 55）。ストレージ装置は、ノード装置 200 からデータを受信して（ステップ S 60）、指定された仮想ブロックに書き込みを行う。

【0059】

以上の処理を繰り返し行うことにより、クライアント 100 は、ストレージ装置 300 の仮想ブロックにデータの書き込みを行うことができる。この際、異常が生じた場合のコアダンプについては、ノード装置へのキャッシュを回避することができる。

【0060】

上述の実施例では、クライアント 100 の異常通知により、ノード装置 200

がキャッシュ可否の属性情報を更新する場合を例示した。これに対し、クライアント100からノード装置200に送信されるデータごとにキャッシュ可否の属性情報を添付するようにしてもよい。例えば、異常制御モジュール141は、コアダンプとして書き込むデータについては、全てキャッシュ不可という属性情報を添付し、ノード装置200は、この属性情報に基づいてキャッシュの可否を判断するようにしてもよい。

【0061】

A6. 変形例;

実施例では、ストレージ装置300からのデータ読み込み時、およびストレージ装置300への書き込み処理時にノード装置200が一切、キャッシュを行わない例を示した。ノード装置200は、キャッシュ可否に関わらず、揮発性メモリ202へのキャッシュを行い、メモリ202からディスク装置210へのデータ移転時にキャッシュの可否を判定するようにしてもよい。かかる制御は、ディスク装置210への書き込み時に、キャッシュ属性管理部234とキャッシュ管理テーブル232を参照することで実現することができる。つまり、キャッシュ管理テーブル232を参照することにより、メモリ202にキャッシュされているデータについて、仮想ボリュームおよびブロック番号を特定することができる。この情報に基づいて、キャッシュ属性管理部234を参照すれば、ディスク装置210へのキャッシュ可否を判定することができる。

【0062】

変形例の制御を実現するための別の方法として、キャッシュ管理テーブル232に、図5の破線で示した「ディスク書込」という情報を設けても良い。「ディスク書込」は、ディスク装置210へのキャッシュ可否を表す属性情報であり、メモリ202へのキャッシュ時に、キャッシュ属性管理部234を参照することで設定することができる。このようにキャッシュ管理テーブル232において、格納されているデータごとに、ディスク書込を保存しておくことにより、ディスク装置210への移転時にキャッシュ属性管理部234を参照する必要がなくなり、キャッシュ可否を簡易に判定することが可能となる。

【0063】

以上で説明した第1実施例によれば、データの読みだし、および書き込み時に、ノード装置200におけるディスク装置210へのキャッシュの可否を切り替えることができる。コアダンプなど、機密情報を含むデータについては、ディスク装置210へのキャッシュを回避することができるため、ノード装置200の盗難による情報漏洩を抑止することができる。ノード装置200においてメモリ202へのキャッシュ自体も禁止する場合には、ノード装置200への不正アクセスによる情報漏洩を抑止することも可能である。

【0064】

B. 第2実施例；

第1実施例では、属性情報に基づいて、ノード装置200におけるキャッシュ可否を切り替える例を示した。第2実施例では、ノード装置が、ディスク装置へのキャッシュ時に、データを暗号化することでセキュリティを確保する例を示す。ノード装置への不正アクセスに対するセキュリティ確保という観点からは、メモリ202へのキャッシュ時にもデータを暗号化してもよい。ただし、本実施例においては、暗号化に要する処理負荷を軽減するため、メモリ202へのキャッシュ時には暗号化を行わず、ディスク装置210へのデータ移転時に暗号化する例を示す。

【0065】

B1. 機能ブロック；

図9は第2実施例としての計算機システムの機能ブロックを示す説明図である。第1実施例と同じ機能ブロックについては、同一の符号で示した。第2実施例では、キャッシュ属性管理部234、キャッシュ属性管理テーブル333に確保されている属性情報は、キャッシュの可否ではなく、暗号化の要否を示す情報となる。第2実施例では、ノード装置200Aに暗号化部235が新たに追加され、ストレージ装置300Aに鍵管理モジュール335が新たに追加される。暗号化部235は、ノード装置200Aにおいて、ディスク装置210へのキャッシュ時に、データの暗号化を行う機能、およびディスク装置210から読み出されたデータを復号する機能を奏する。鍵管理モジュール335は、ノード装置200Aに対して、暗号化に使用する鍵情報を管理・提供する機能を奏する。

【0066】

B2. データ読出処理；

図10は第2実施例におけるデータ読出処理のフローチャートである。左側にクライアント100の処理、中央にノード装置200Aの処理、右側にストレージ装置300Aの処理を示した。第2実施例では、ノード装置200Aがネットワークに接続され、起動すると、ストレージ装置300Aは、ノード装置200Aに対して鍵情報の送信を行う（ステップS31）。ノード装置200Aは、この鍵情報を暗号化部235に記憶する（ステップS70）。

【0067】

本実施例においては、鍵情報は、ノード装置ごとに固有の情報として設定される。ストレージ装置300は、ネットワークに接続されている複数のノード装置について、提供した鍵情報を一元的に管理し、新たに接続されたノード装置に対しては、これらのいずれとも異なる鍵情報を生成する。鍵情報は、例えば、シリアル番号などの規則的なものであってもよいし、乱数を用いた不規則な情報であってもよい。各ノード装置に一意的な鍵情報を提供することにより、いわゆる悪意の第三者がノード装置になりすまして鍵情報を取得した場合でも、各ノード装置にキャッシュされたデータのセキュリティを確保することができる利点がある。

【0068】

クライアント100がノード装置200Aに対して、データの読みだし要求を送信すると（ステップS12）、ノード装置200Aは、要求されたデータが、キャッシュ済みであるか否かを判断する（ステップS71）。キャッシュ済みである場合には、メモリ202またはディスク装置210からデータを読みだし（ステップS72）、暗号化されている場合には、復号して（ステップS73、S74）、クライアント100に返信する（ステップS81、ステップS13）。

【0069】

要求されたデータがキャッシュされていない場合（ステップS71）、ノード装置200Aは、ストレージ装置300Aにデータを要求する（ステップS75）。ストレージ装置300Aは、この要求に応じて、データをノード装置200Aに返信する（ステップS32）。これと併せて、ストレージ装置300は、要

求されたデータについて暗号化の要否を表す属性情報をノード装置 200A に通知する (ステップ S32)。この属性情報は、要求されたデータのヘッダ等に埋め込んだ形で通知してもよいし、要求されたデータと分けて通知してもよい。

【0070】

ノード装置 200A は、ストレージ装置 300A からデータおよび属性情報を受信すると、データのキャッシュ先を判断する (ステップ S76)。メモリ 202 にキャッシュ可能な領域があいている場合には、メモリ 202 へのキャッシュを行う。キャッシュ可能な領域が空いていない場合には、ディスク装置 210 へのキャッシュが必要と判断する。この判断は、ストレージ装置 300A からのデータ受信時に行っても良いし、いわゆる LRU (Least Recently Used) に基づいて処理してもよい。LRU とは、メモリ 202 にキャッシュされているデータのうち、アクセス頻度の低いデータを優先的にディスク装置 210 に移転する方法である。LRU で処理する場合には、ストレージ装置 300A から受信したデータは、無条件にメモリ 202 にキャッシュされ、代わりに、従前にメモリ 202 に蓄積されていたデータのいずれかがディスク装置 210 に移転されることになる。

【0071】

ノード装置 200A は、ディスク装置 210 にキャッシュすべきと判断されたデータについては (ステップ S76)、属性情報に基づいて暗号化の要否を判断する (ステップ S77)。暗号化が要求されている場合には、暗号化を行い (ステップ S78)、キャッシュする (ステップ S79)。暗号化が要求されていないデータ、またはメモリ 202 にキャッシュすべきデータについては (ステップ S76)、暗号化を行わずにキャッシュする。ノード装置 200A は、これらのキャッシュの格納先に応じて、キャッシュ管理テーブル 232 を更新し (ステップ S80)、データをクライアント 100 に返信する (ステップ S25)。クライアントはこうしてノード装置 200 から送信されたデータを受信する (ステップ S13)。

【0072】

以上の処理を繰り返し行うことにより、クライアント 100 は、必要なファイ

ルをストレージ装置 300A から読み出すことができる。この際、機密情報を含むデータについては、属性情報の設定により、ディスク装置 210 へのキャッシュ時に暗号化することができるため、セキュリティを確保することができる。第 2 実施例においても第 1 実施例と同様、ノード装置 200A が、暗号要否の属性情報をストレージ装置 300A から、予めまとめて取得し、管理するようにしてもよい。

【0073】

B3. データ書込処理；

図 11 は第 2 実施例におけるデータ書込処理のフローチャートである。左側にクライアント 100 の処理、中央にノード装置 200A の処理、右側にストレージ装置 300A の処理を、コアダンプの書き込みを例にとって示した。

【0074】

読みだし処理と同じく、ノード装置 200A は起動時に、ストレージ装置から鍵情報を受信し、記憶している（ステップ S60、S31）。また、第 1 実施例と同様、クライアント 100 の異常制御モジュール 141 は、ソフトウェア上の動作異常を検出すると（ステップ S40）、異常の発生を IP アドレスとともにノード装置 200 に通知する（ステップ S40）。ノード装置 200 は、この情報に基づいてキャッシュ属性管理部 234 における属性情報を暗号化要という内容に更新する（ステップ S50）。

【0075】

この状態で、クライアント 100 がコアダンプを書き込むための要求をノード装置 200A に送信すると（ステップ S42）、ノード装置 200A は、読みだし処理と同様、受信したデータのキャッシュ先を判断する。ノード装置 200A は、ディスク装置 210 にキャッシュすべきと判断されたデータについて、属性情報で指定された暗号化の要否に基づいて暗号化を行った上で、キャッシュする（ステップ S90～S93）。暗号化が要求されていないデータ、またはメモリ 202 にキャッシュすべきデータについては（ステップ S90）、暗号化を行わずにキャッシュする（ステップ S94）。ノード装置 200A は、これらのキャッシュの格納先に応じて、キャッシュ管理テーブル 232 を更新し（ステップ S

94)、データをストレージ装置300Aに送信する(ステップS95)。ストレージ装置300Aは、ノード装置200Aからデータを受信して(ステップS60)、指定された仮想ブロックに書き込みを行う。以上の処理を繰り返し行うことにより、クライアント100は、ストレージ装置300Aの仮想ブロックにデータの書き込みを行うことができる。

【0076】

第2実施例によれば、ノード装置200Aにおいて、ディスク装置への書き込み時に暗号化することにより、盗難によるデータの漏洩を抑止することができる。メモリ202への書き込み時には暗号化を行わないため、頻繁に使用されるデータについては、メモリ202にキャッシュしておくことにより、読みだし、書き込みの速度を確保することもできる。

【0077】

第2実施例では、鍵情報をストレージ装置300Aから提供したが、鍵情報は厳密に一意である必要はなく、ノード装置200Aで生成しても構わない。第2実施例では、属性情報によって暗号化の可否を制御する場合を例示したが、ディスク装置210へのキャッシュ時には必ず暗号化するようにしてもよい。

【0078】

C. 第3実施例；

第3実施例では、キャッシュ可とされているデータについては、クライアントとストレージ装置間でノード装置を介してデータの授受を行い、キャッシュ不可とされているデータについては、両者間で直接データの授受を行う場合を例示する。

【0079】

C1. 機能ブロック；

図12は第3実施例としての計算機システムの機能ブロックを示す説明図である。第1実施例と同じ機能ブロックについては、同一の符号で示した。第3実施例では、クライアント100に提供されるオペレーティングシステム140Bの内容が、第1実施例と相違する。オペレーティングシステム140Bは、SCSIコマンドの送出先としてノード装置200Bのアドレス(以下、「ノードアド

レス」と称する) 142を保持している。オペレーティングシステム140Bには、動作異常時にコアダンプの書き込みなどを行う異常制御モジュール141Bが供えられており、この異常制御モジュール141Bは、コアダンプの書き込み先としてストレージ装置300Bのアドレス(以下、「ストレージアドレス」と称する) 143を保持している。これらの機能ブロックにより、クライアントは、異常時と正常時でデータの書き込み先を切り替えることができる。

【0080】

ストレージ装置300Bにおいては、アクセス制御モジュール330Bが、キャッシュ属性に基づいてデータの送信先を切り替える。キャッシュ可とされているデータについてはノード装置に送信し、キャッシュ不可とされているデータはクライアント100に直接送信する。第3実施例では、キャッシュ不可とされているデータについては、クライアント100とストレージ装置300Bの間で直接送受信されるから、ノード装置200Bにおいて、キャッシュ制御モジュール230Bは、クライアント100およびストレージ装置300Bから受信したデータについては、キャッシュ可否を考慮することなく、データをキャッシュする。

【0081】

C2. データ読出処理;

図13は第3実施例におけるデータ読み出し処理のフローチャートである。リモートブート時を例にとり、左側にクライアント100の処理、中央にノード装置200Bの処理、右側にストレージ装置300Bの処理を示した。

【0082】

ユーザが計算機の電源をオンにすると、第1実施例と同様の手順でログインおよびブートプログラムの起動が行われる(ステップS10、S30、S11)。クライアント100は、仮想ボリュームを指定して、データの読み出し要求をノード装置200Bに対して送信する(ステップS12)。

【0083】

ノード装置200Bは、要求されたデータがキャッシュ済みの場合は(ステップS100)、キャッシュブロックから該当するデータを読み出し(ステップS

101)、クライアント100に返信する(ステップS105)。クライアント100は、このデータを受信して起動処理を継続する(ステップS13)。

【0084】

要求されたデータが未キャッシュの場合(ステップS100)、ノード装置200Bはストレージ装置300Bにデータの読み出しを要求する(ステップS102)。ストレージ装置300Bは、キャッシュ属性管理テーブル333において、要求されたデータがキャッシュ可と設定されている場合には(ステップS110)、データをノード装置に返信する(ステップS111)。ノード装置200Bは、このデータをキャッシュし(ステップS103)、キャッシュ管理テーブル232を更新して(ステップS24)、データをクライアントに返信する(ステップS105)。

【0085】

要求されたデータがキャッシュ不可の場合には(ステップS110)、ストレージ装置300Bは、データをクライアントに直接、返信する(ステップS112)。クライアントはこうしてノード装置200から送信されたデータを受信する(ステップS13)。ストレージ装置300Bは、ノード装置200Bに対して、クライアントへの返信を行った旨の通知を行っても良い。ノード装置200Bが、データの要求(ステップS102)へのレスポンスを待っている場合には、この通知により、レスポンス待ちの状態から解放することができる。

【0086】

上記実施例において、ストレージ装置300Bからクライアントに直接、データを送信する場合、クライアントのアドレスは種々の方法で特定することができる。例えば、ストレージ装置300Bは、最初のログイン時にユーザ管理テーブル334に記録されているIPアドレスを利用してもよい。別の例として、データの要求時に、ノード装置200Bからストレージ装置300Bに対して、クライアント100のアドレス情報を通知するようにしてもよい。

【0087】

C3. データ書込処理;

図14はデータ書込処理のフローチャートである。左側にクライアント100

の処理、中央にノード装置 200B の処理、右側にストレージ装置 300B の処理を、コアダンプの書き込みを例にとって示した。クライアント 100 は、異常が生じていない場合には（ステップ S120）、ノード装置 200 に対してデータの書き込み要求を送信する（ステップ S121）。ノード装置 200B は、この要求を受信すると、データをキャッシュし（ステップ S13）、キャッシュ管理テーブル 232 を更新する（ステップ S131）。その後、キャッシュされている書き込みデータが所定量に達したなど、予め設定されたタイミングで、データをストレージ装置 300B に送信する（ステップ S132）。ストレージ装置 300B は、このデータを受信して、指定された仮想ブロックにデータの書き込みを行う（ステップ S140）。

【0088】

一方、異常制御モジュール 141 が、ソフトウェア上の動作異常を検出すると（ステップ S120）、データの送信先をノード装置 200 のアドレスからストレージ装置に切り替えることにより、ストレージ装置 300B に対して直接、データの書き込み要求を送信する（ステップ S122）。ストレージ装置 300B は、このデータを受信して、指定された仮想ブロックにデータの書き込みを行う（ステップ S140）。

【0089】

第3実施例によれば、キャッシュの可否によって、ノード装置 200B を迂回してデータを送受信することができる。従って、ノード装置の盗難や不正アクセスに対し、データのセキュリティを確保することができる。

【0090】

以上、本発明の種々の実施例について説明したが、本発明はこれらの実施例に限定されず、その趣旨を逸脱しない範囲で種々の構成を採ることができることはいうまでもない。

【0091】

【発明の効果】

本発明によれば、ノード装置へのキャッシュ自体の回避、キャッシュ時の暗号化、ノード装置を迂回したデータ授受などの方法により、機密情報を含んだデー

タ等がノード装置のディスク装置に、原状でキャッシュされることを回避することができる。従って、ノード装置の盗難、不正アクセスに対して、これらのデータのセキュリティを確保することができる。

【図面の簡単な説明】

【図 1】 第 1 実施例としての計算機システムの構成を示す説明図である。

【図 2】 計算機システムを構成する各装置内の機能ブロックを示す説明図である。

【図 3】 ユーザ管理テーブル 334 の構造を示す説明図である。

【図 4】 物理ブロック管理テーブル 332 の構造を示す説明図である。

【図 5】 キャッシュ管理テーブル 232 の構造を示す説明図である。

【図 6】 キャッシュ属性管理テーブル 333 の構造を示す説明図である。

【図 7】 データ読み出し処理のフローチャートである。

【図 8】 データ書込処理のフローチャートである。

【図 9】 第 2 実施例としての計算機システムの機能ブロックを示す説明図である。

【図 10】 第 2 実施例におけるデータ読出処理のフローチャートである。

【図 11】 第 2 実施例におけるデータ書込処理のフローチャートである。

【図 12】 第 3 実施例としての計算機システムの機能ブロックを示す説明図である。

【図 13】 第 3 実施例におけるデータ読み出し処理のフローチャートである。

【図 14】 データ書込処理のフローチャートである。

【符号の説明】

100…クライアント

101、201、301…CPU

102、202、302…メモリ (RAM)

103、203、303…ROM

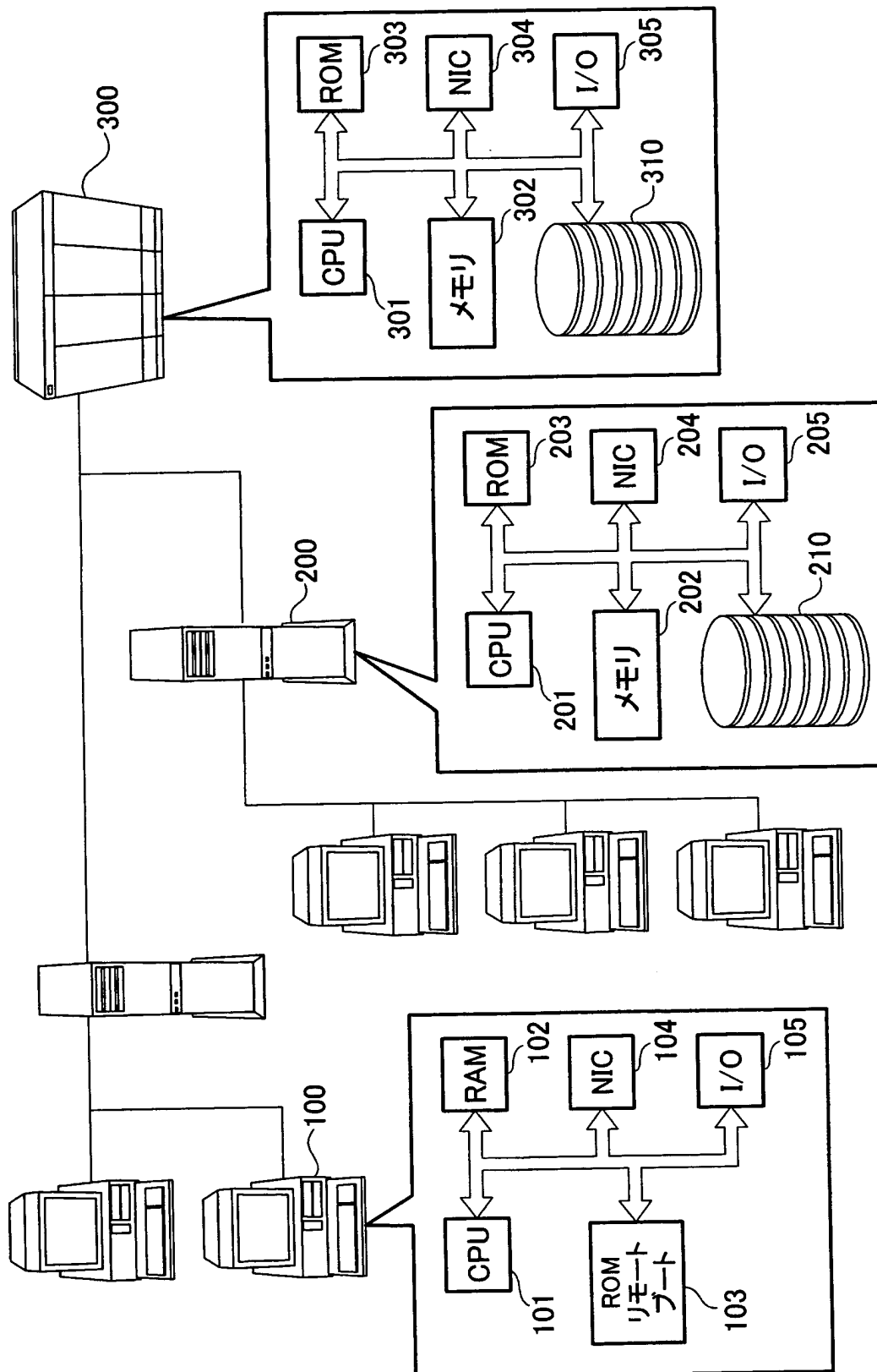
104、204、304…NIC

105、205、305…I/O

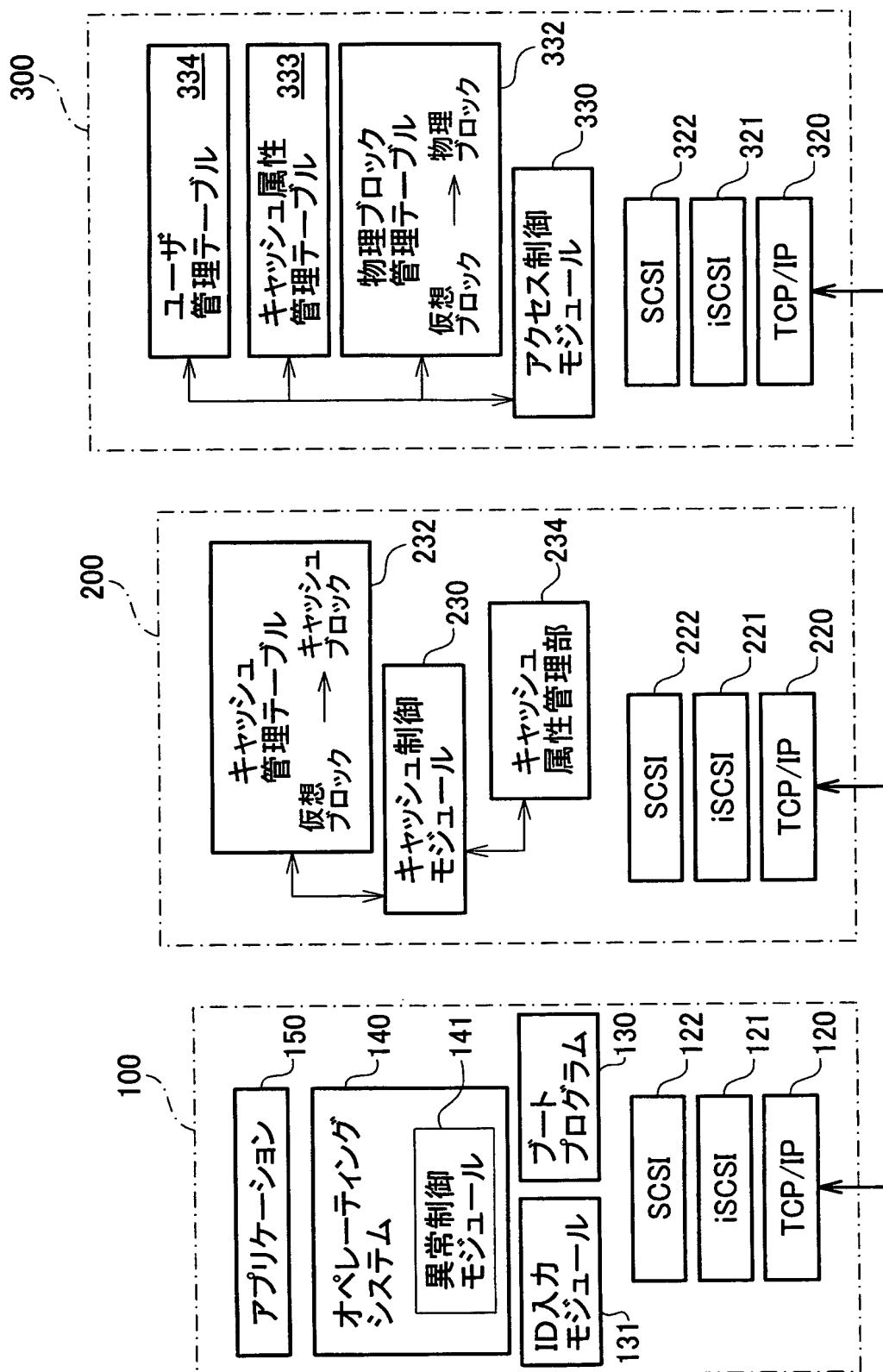
120、220、320…TCP/IPモジュール
121、221、321…iSCSIモジュール
122、222、322…SCSIモジュール
130…ブートプログラム
131…ID入力モジュール
140、140B…オペレーティングシステム
141、141B…異常制御モジュール
142…ノードアドレス
143…ストレージアドレス
150…アプリケーション
200、200A、200B…ノード装置
210、310…ディスク装置
230、230B…キャッシュ制御モジュール
232…キャッシュ管理テーブル
234…キャッシュ属性管理部
235…暗号化部
300、300A、300B…ストレージ装置
330、330B…アクセス制御モジュール
332…物理ブロック管理テーブル
333…キャッシュ属性管理テーブル
334…ユーザ管理テーブル

【書類名】 図面

【図 1】



【図 2】

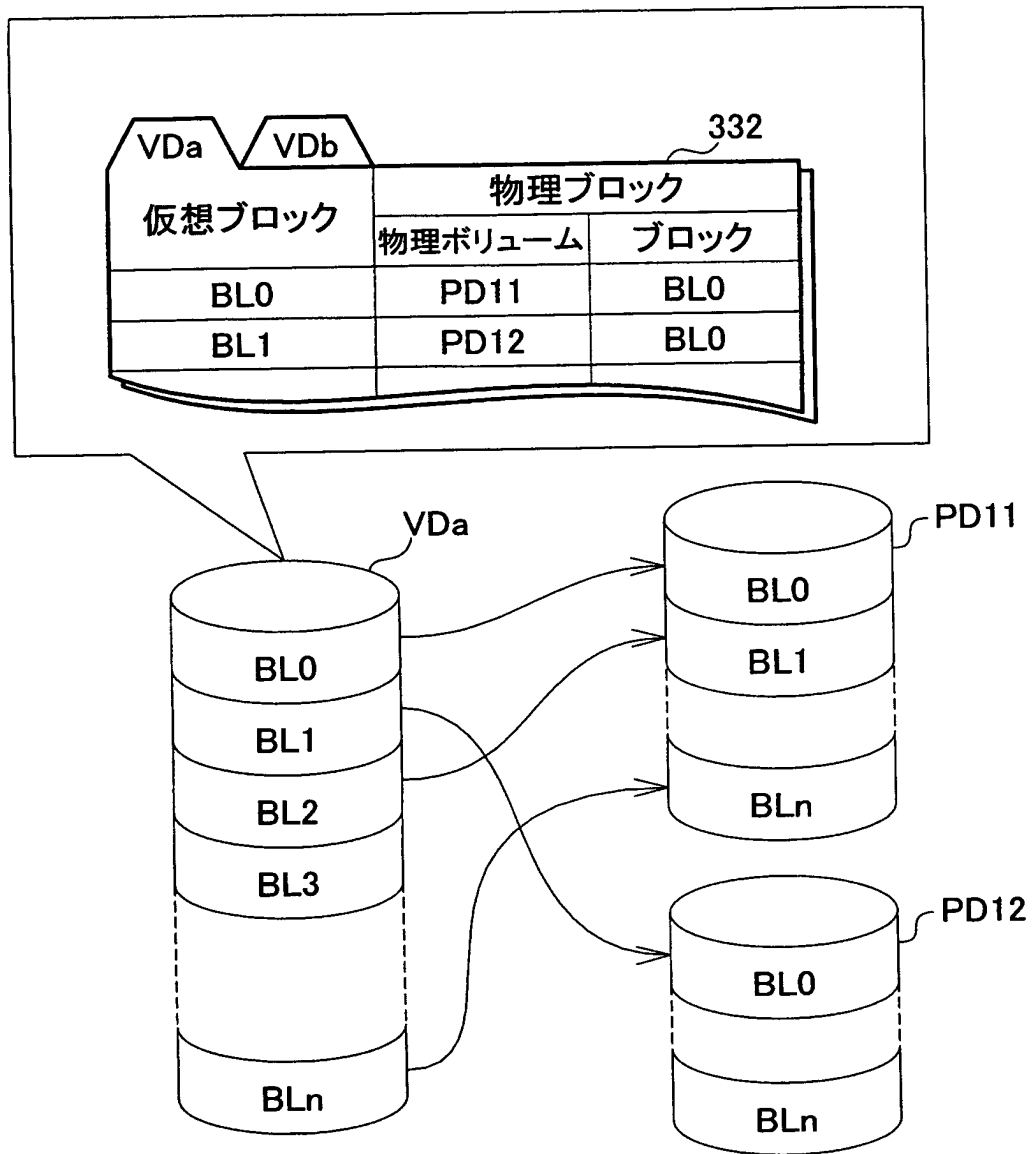


【図 3】

334

ユーザ名	パスワード	仮想ボリューム	属性	IPアドレス
ユーザA	*****	VDa	専有	IPa
ユーザB	*****	VDb	専有	
ユーザC	*****	VDC	共有	

【図 4】



【図 5】

232

仮想ブロック		キャッシュブロック		ディスク書込
仮想ボリューム	ブロック	格納場所	ブロック	
VDa	BL0	ディスク	BL0	可
VDb	BL0	メモリ	BL1	可
VDa	BL1	メモリ	BL0	不可
VDb	BL2	ディスク	BL3	可

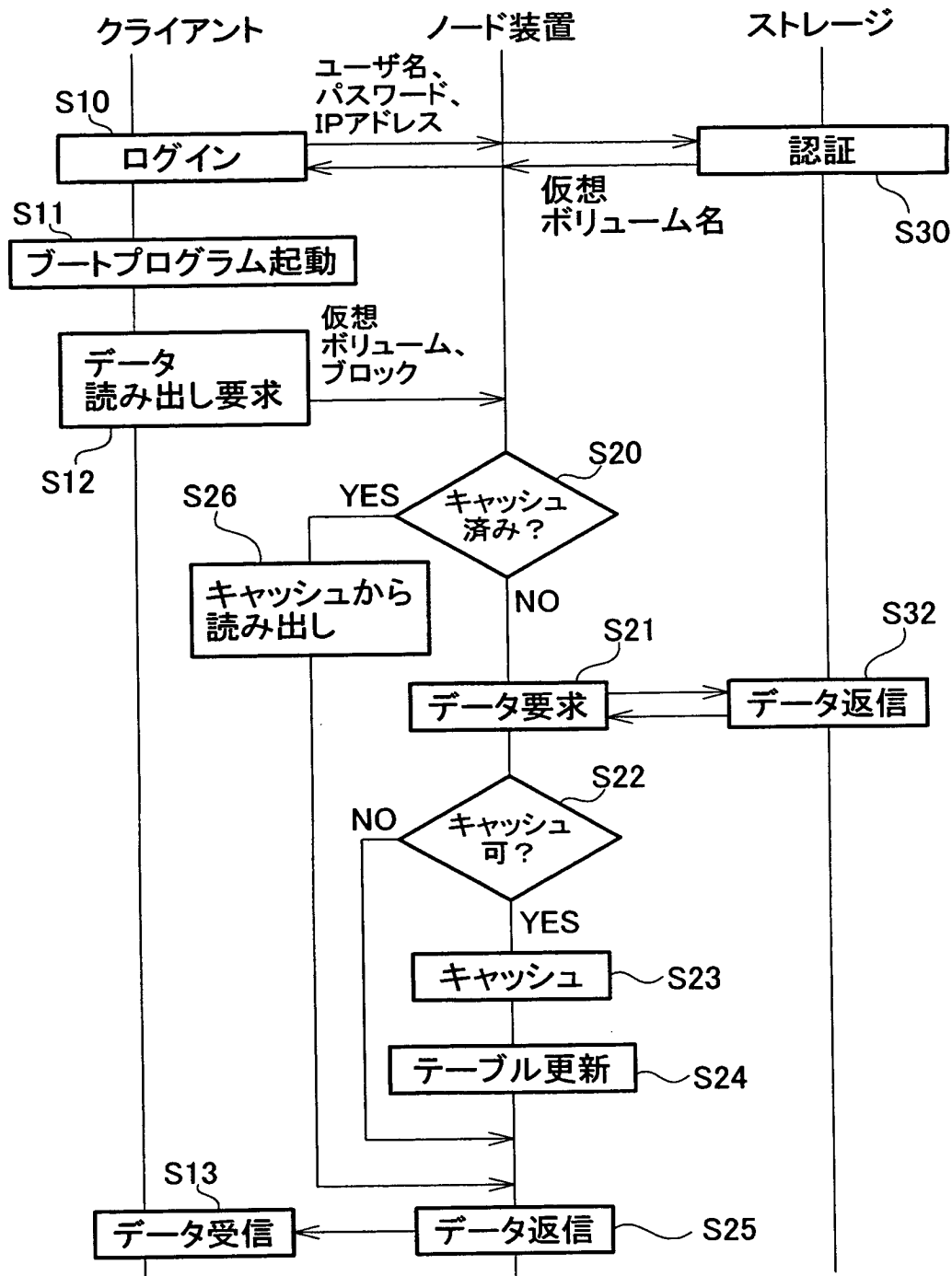
【図 6】

仮想ボリューム	読み出し		書き込み	
	物理ボリューム	キャッシュ	物理ボリューム	キャッシュ
VDa	PD11	可	PD11	不可
	PD12	不可	PD12	不可
	・ ・		・ ・	・ ・
VDb	PD21	可	PD21	可
	PD12	不可	PD12	不可
	・ ・		・ ・	

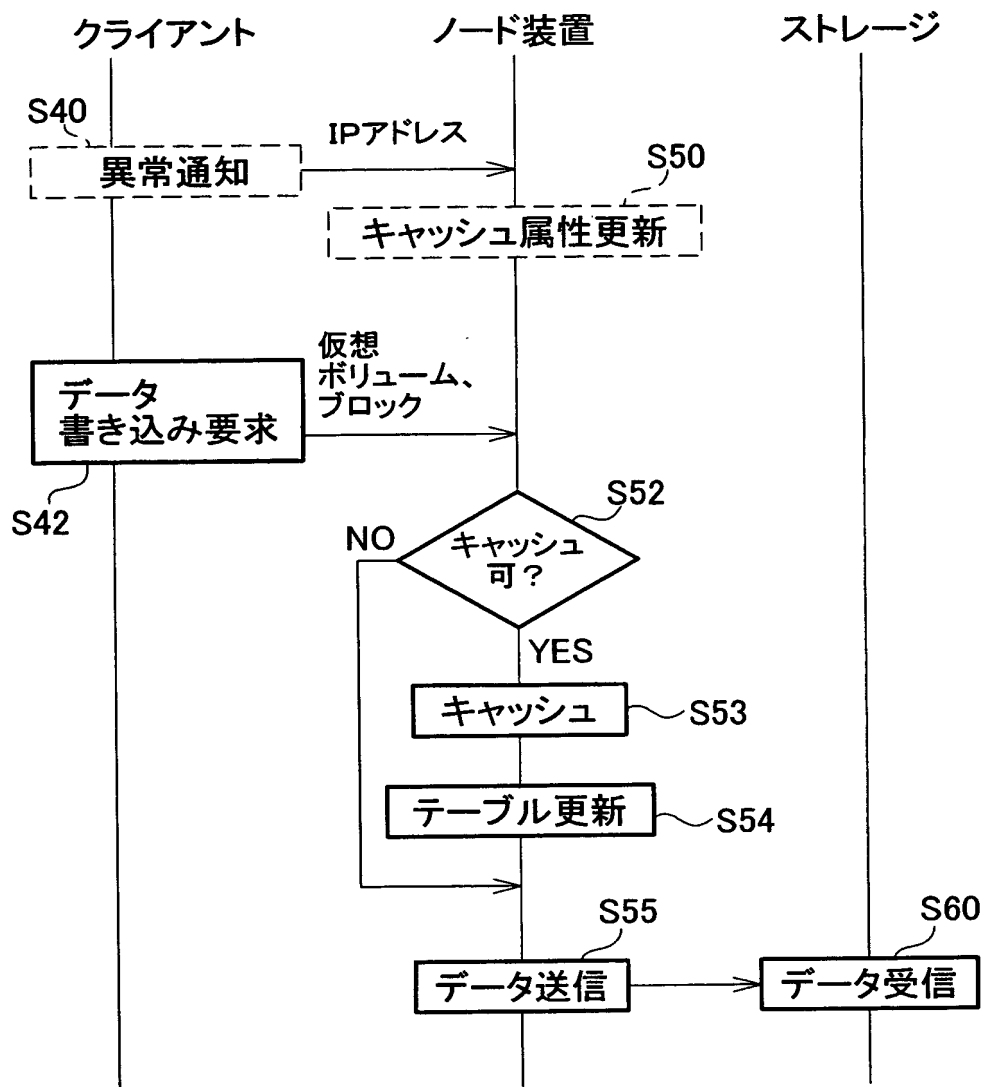
234

Da

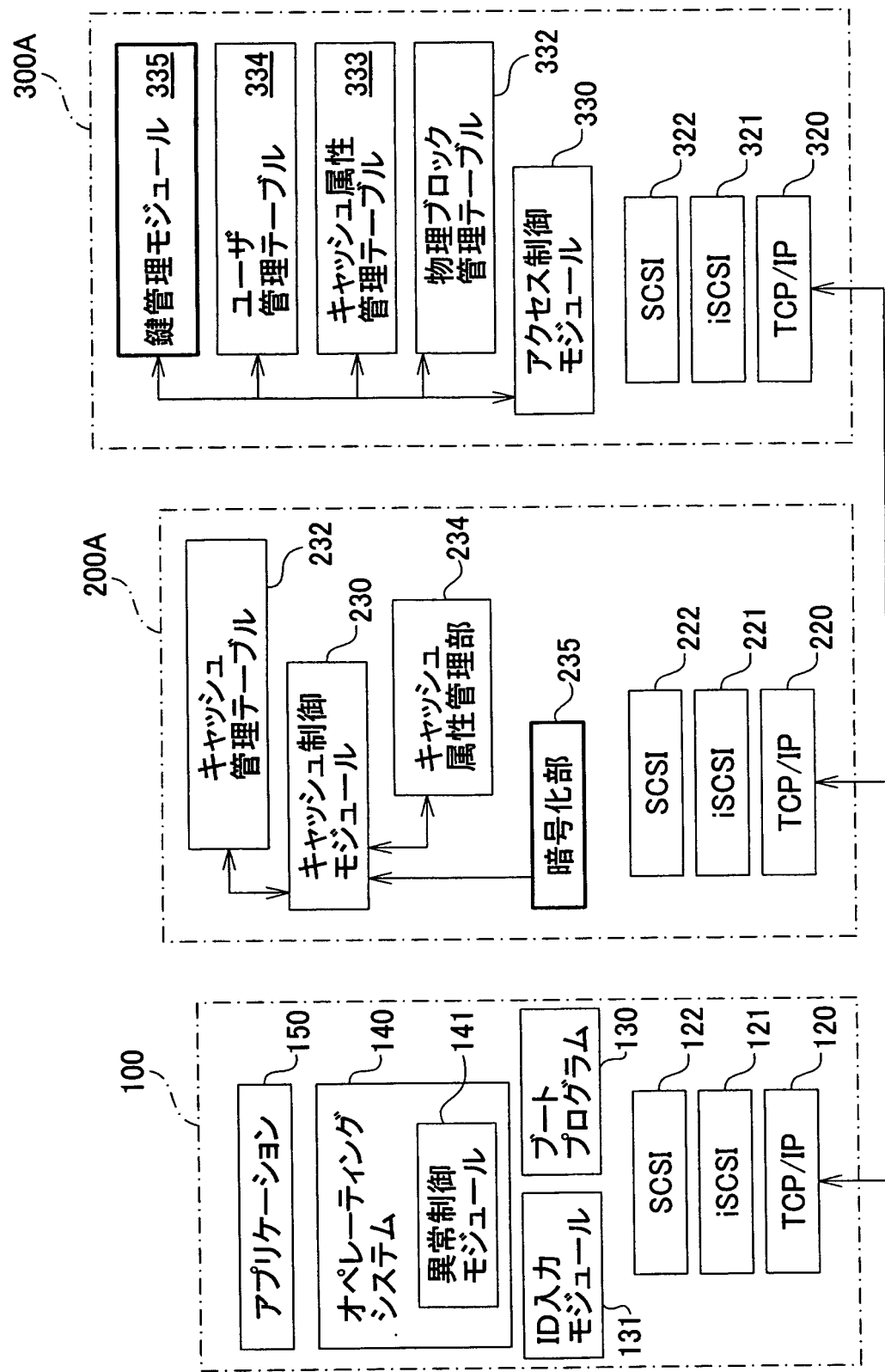
【図 7】



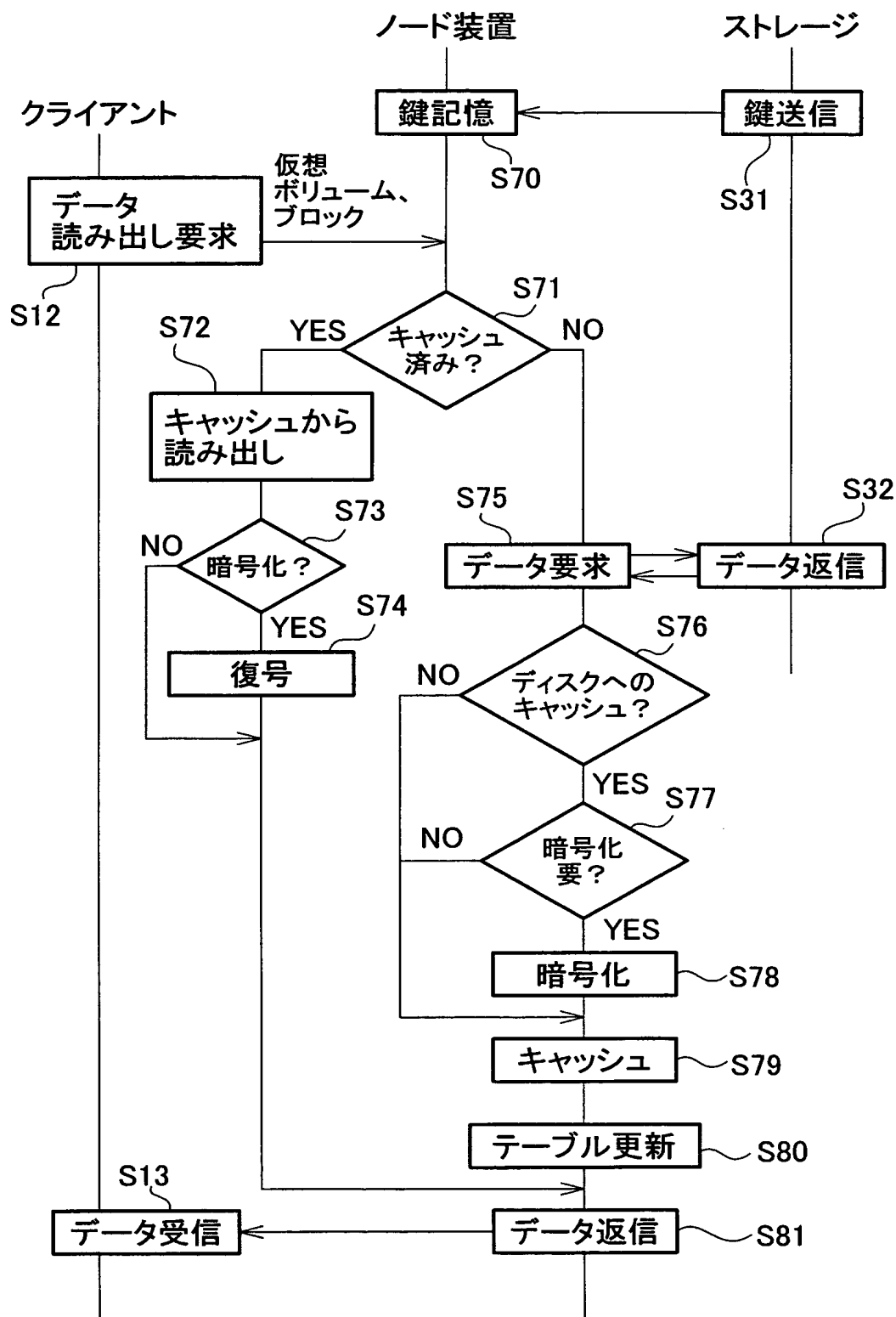
【図 8】



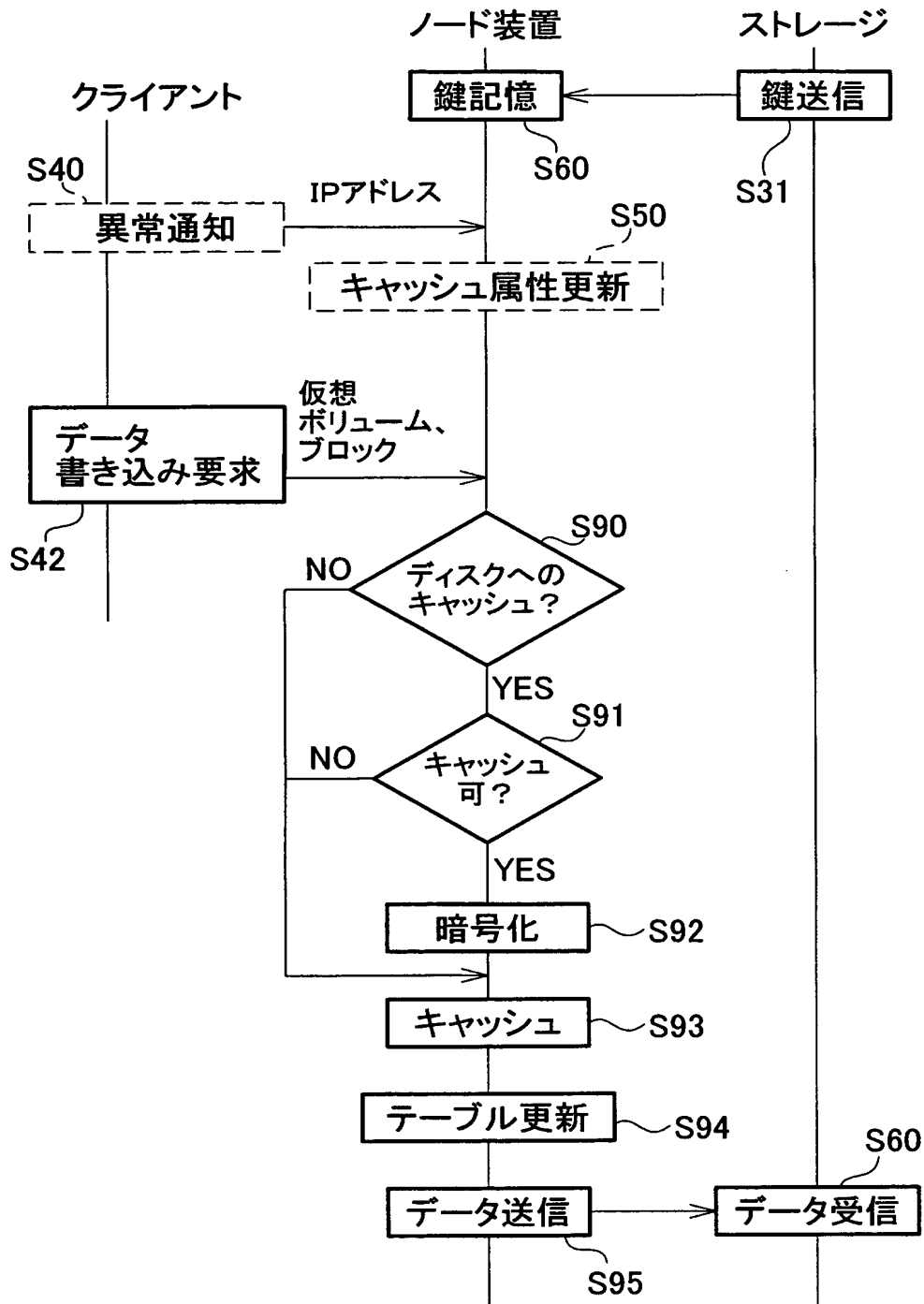
【図 9】



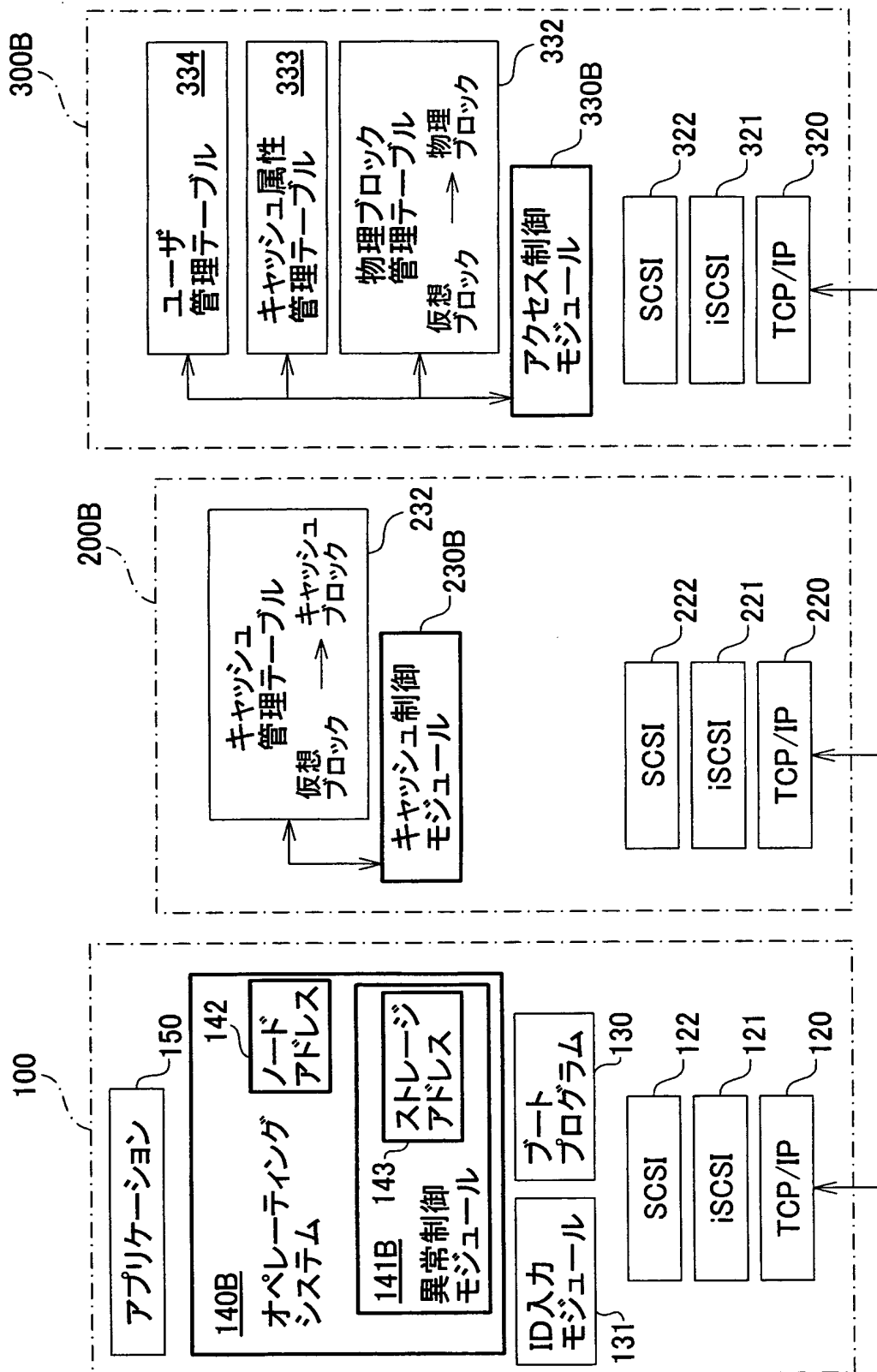
【図10】



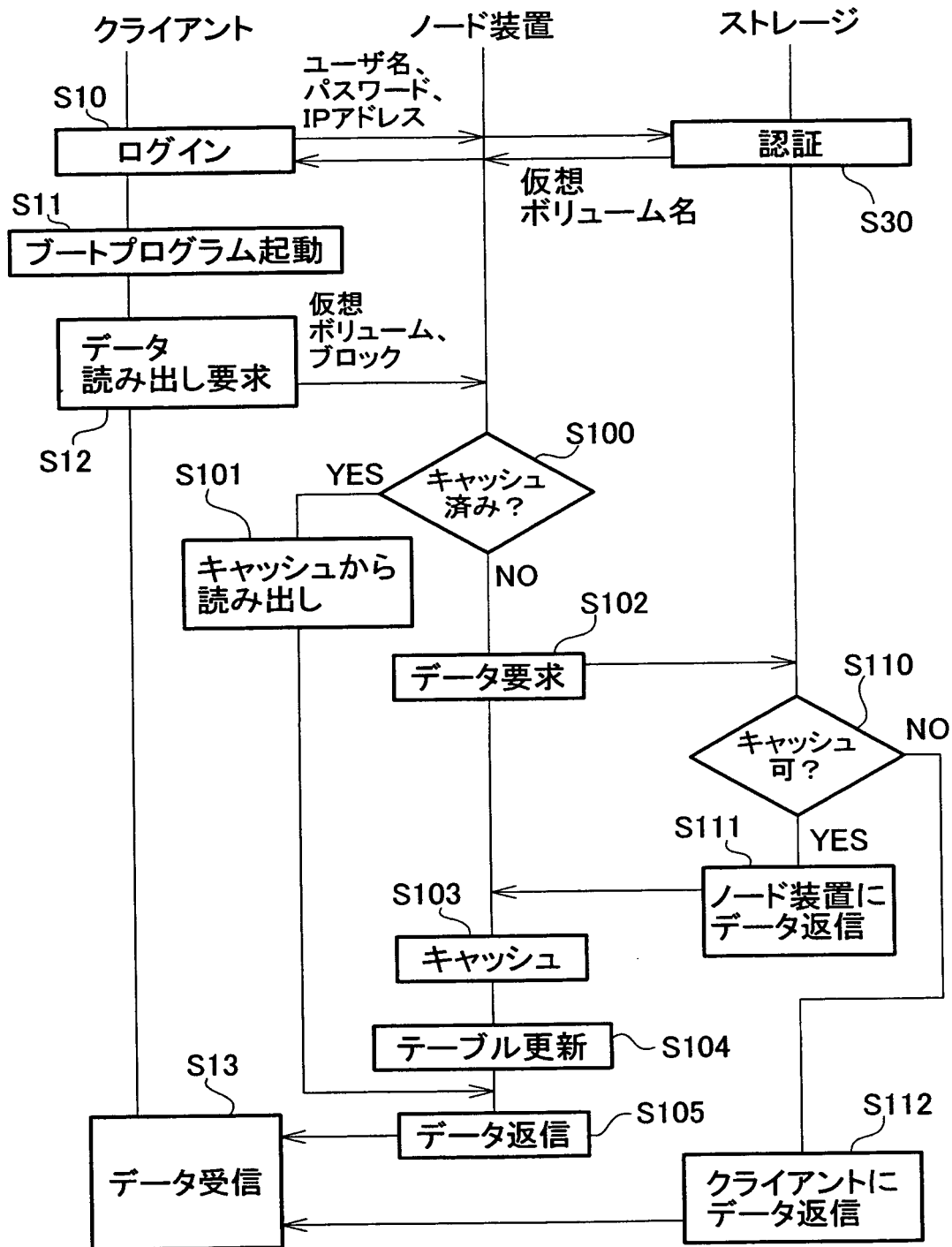
【図 11】



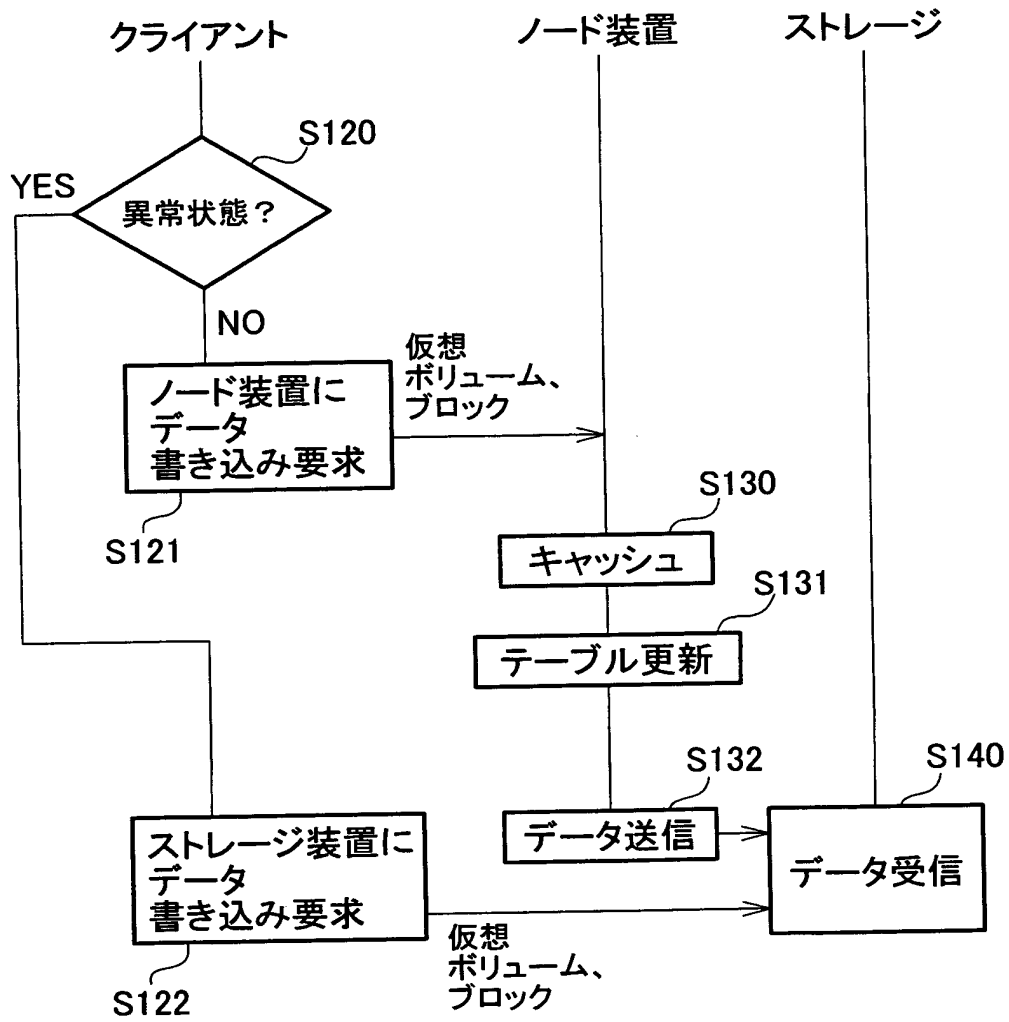
【図 12】



【図 13】



【図 14】



【書類名】 要約書

【要約】

【課題】 クライアント、ノード装置、ストレージ装置からなる計算機システムにおいて、ノード装置の盗難、不正アクセスに対するデータのセキュリティを向上する。

【解決手段】 クライアントおよびストレージ装置間で、ノード装置を介してデータを授受可能な計算機システムにおいて、ノード装置におけるキャッシュを次の方法で制御する。第1の制御では、ストレージ装置に格納されているデータについてキャッシュ可否の属性を付与しておき、ノード装置は、キャッシュ不可とされているデータについては、キャッシュを行わずにクライアントへの中継を行う。第2の制御では、ノード装置は、ディスクへのキャッシュ時にデータの暗号化を行う。第3の制御では、キャッシュ不可のデータについては、ノード装置を介さずに直接送受信する。これらの制御により、ノード装置へのキャッシュを制限でき、セキュリティを確保することができる。

【選択図】 図1

特願 2003-075181

出願人履歴情報

識別番号

[000005108]

1. 変更年月日

1990年 8月31日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台4丁目6番地

氏 名

株式会社日立製作所